GENERATING AND TRANSMITTING INVOCATION REQUEST TO APPROPRIATE THIRD-PARTY AGENT

BACKGROUND

An automated assistant (also known as "personal assistant", "mobile assistant", or "chat bot") may be interacted with by a user via a variety of client devices, such as smart phones, tablet computers, wearable devices, automobile systems, standalone personal assistant devices, and so forth. An automated assistant receives input from the user (e.g., typed and/or spoken natural language input) and responds with responsive content (e.g., visual and/or audible natural language output). An automated assistant interacted with via a client device may be implemented via the client device itself and/or via one or more remote computing devices that are in network communication with the client device (e.g., computing device(s) in "the cloud").

SUMMARY

This specification is directed generally to selectively invoking third-party (3P) agents. Some implementations are directed to selective invocation of a particular 3P agent by an automated assistant to achieve an intended action determined by the automated assistant during a dynamic dialog between the automated assistant and a user. In some of those implementations, the particular 3P agent is invoked with value(s) for parameter(s) that are determined during the dynamic dialog; and/or the particular 3P agent is selected, from a plurality of candidate 3P agents, for invocation based on the determined value(s) for the parameter(s) and/or based on other criteria. In some versions of those implementations, the automated assistant invokes the particular 3P agent by transmitting, to the particular 3P agent, a 3P invocation request that includes the determined value(s) for the parameter(s). For example, the automated assistant may transmit the 3P invocation request to the 3P agent over one or more networks and may interface with the 3P agent utilizing an application programming interface (API).

When the particular 3P agent is invoked, the particular 3P agent generates 3P responsive content that can be used to generate output for inclusion in the dynamic dialog. When the invocation request includes the determined value(s) for the parameter(s), the particular 3P agent may take those value(s) into account in generating the 3P responsive content. This may enable the intended action to be achieved via the particular 3P agent more quickly, which may conserve various computational resources (e.g., network traffic consumed by additional "turns" that might be necessary absent the provision of the value(s)).

Moreover, when the determined value(s) for the parameter(s) are utilized to select the particular 3P agent, it may ensure that the selected particular 3P agent is able to perform the intended action based on the determined values. For example, determined value(s) for parameter(s) may be compared to corresponding value(s) for the parameter(s) defined for each of a plurality of candidate 3P agents that can perform the intended action—and only the 3P agent(s) whose value(s) correspond to the determined value(s) can be considered for selection. This may mitigate the risk that the particular 3P agent selected for invocation is unable to perform the intended action, which may conserve various computational resources. For example, it may conserve network and/or processor resources that may otherwise be consumed by an initial failed attempt to utilize a 3P agent to perform the intended action, that is then followed by invoking an alternative 3P agent in another attempt to perform the intended action.

In some situations, in response to invocation of the particular 3P agent, value(s) for parameter(s) that are provided with an invocation request may enable the particular 3P agent to achieve the intended action without engaging in further dialog with the user. In those situations, the 3P agent may provide responsive content that indicates the intended action has been achieved and/or that indicates additional information related to achievement of the intended action.

In some situations, in response to invocation of the particular 3P agent, the dynamic dialog may be at least temporarily transferred (actually or effectively) to the particular 3P agent, such that the 3P agent at least temporarily "steers" the substance of the dynamic dialog. For example, output that is based on the 3P responsive content may be provided to the user in furtherance of the dialog, and further user input received in response to the 3P responsive content. The further user input (or a conversion thereof) may be provided to the particular 3P agent. The particular 3P agent may utilize its own semantic engines and/or other components in generating further responsive content that can be used to generate further output for providing in furtherance of the dynamic dialog. This general process may continue until, for example, the 3P agent provides responsive content that terminates the 3P agent dialog (e.g., an answer or resolution instead of a prompt), additional user interface input of the user terminates the 3P agent dialog (e.g., instead invokes a response from the automated assistant or another 3P agent), etc.

As described herein, in some situations the automated assistant may still serve as an intermediary when the dialog is effectively transferred to the 3P agent. For example, in serving as an intermediary where natural language input of a user is voice input, the automated assistant may convert that voice input to text, provide the text (and optionally annotations of the text) to the 3P agent, receive 3P responsive content from the 3P agent, and provide output that is based on the 3P responsive content for presentation to the user. Also, for example, in serving as an intermediary, the automated assistant may analyze user input and/or 3P responsive content to determine if dialog with the 3P agent should be terminated, transferred to an alternate 3P agent, etc. As also described herein, in some situations the dialog may be actually transferred to the 3P agent (without the automated assistant serving as an intermediary once transferred), and transferred back to the automated assistant upon occurrence of one or more conditions such as termination by the 3P agent (e.g., in response to completion of an intended action via the 3P agent).

Implementations described herein enable an automated assistant to select an appropriate 3P agent based on dialog with a user, and invoke the 3P agent to achieve an intended action of the user that is indicated by the user in the dialog. These implementations may enable a user to engage a 3P agent via interaction with an automated assistant, without necessitating the user know "hot words" to explicitly trigger the 3P agent and/or without necessitating the user even initially know the 3P agent exists. Moreover, implementations enable the automated assistant to determine, based on the dialog with the user, values for various parameters (associated with the intended action), and to pass those parameters to the 3P agent in conjunction with invoking the 3P agent. Moreover, these implementations may enable the user to utilize a common automated assistant interface (e.g., an audible/voice-based interface and/or graphical interface)